

08 OCT 2004

## RAW SEQUENCE LISTING

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) no errors detected.

Application Serial Number: 10/510,941  
Source: PCT  
Date Processed by STIC: 10/22/04

# ***ENTERED***

08 OCT 2004

**CRF Errors Edited by the STIC Systems  
Branch**

Serial Number: 101510,941

CRF Edit Date: 10-22-04  
Edited by: KL

Realigned nucleic acid/amino acid numbers/text in cases where the sequence text "wrapped" to the next line

Corrected the SEQ ID NO. Sequence numbers edited were:

Inserted or corrected a nucleic number at the end of a nucleic line. SEQ ID NO's edited:

Deleted: ✓ invalid beginning/end-of-file text ; \_\_ page numbers

Inserted mandatory headings/numeric identifiers, specifically:

Moved responses to same line as heading/numeric identifier, specifically:

Other:

SEARCHED PCT 070 08 OCT 2004



PCT

**RAW SEQUENCE LISTING**  
**PATENT APPLICATION: US/10/510,941**

**DATE: 10/22/2004**  
**TIME: 09:18:29**

**Input Set : A:\pto.kd.txt**  
**Output Set: N:\CRF4\10222004\J510941.raw**

3 <110> APPLICANT: Jorgensen, Steen Troels  
 4 Rasmussen, Michael Dolberg  
 5 Andersen, Jens Tonne  
 6 Olesen, Peter Bjarke  
 7 Clausen, Ib Groth  
 9 <120> TITLE OF INVENTION: Improved Bacillus Host Cell  
 11 <130> FILE REFERENCE: 10297.204-US  
 C--> 13 <140> CURRENT APPLICATION NUMBER: US/10/510,941  
 C--> 13 <141> CURRENT FILING DATE: 2004-10-08  
 13 <160> NUMBER OF SEQ ID NOS: 22  
 15 <170> SOFTWARE: PatentIn version 3.3  
 17 <210> SEQ ID NO: 1  
 18 <211> LENGTH: 1602  
 19 <212> TYPE: DNA  
 20 <213> ORGANISM: Bacillus licheniformis  
 23 <220> FEATURE:  
 24 <221> NAME/KEY: CDS  
 25 <222> LOCATION: (501)..(1421)  
 27 <400> SEQUENCE: 1  
 28 aacagccaga aaggagacga agaacgtctc tatttaacag agcttgatat gttctgcgtc 60  
 30 catttcgggt caaggatata ttcaactatta tgaaaccaat tttccgatt tccgtataaa 120  
 32 caaagcgtgg aacctgttaa aggacaagaa aaatcgacag ggatataaaa atcggcaagc 180  
 34 ttcccaaagg gaaccggggc aggctgaagc tagcgtcac attggcgagc gatgcagcgg 240  
 36 tcgtgcgtt tcatgaacca tccggctctcc gatggtgagg gagtcgacag tcaaaagcct 300  
 38 tcttacatat ttggagctt acaaggcagac aatcgccatc gccacctatg aaattgaaca 360  
 40 aatattggat gtcgtcttaa atggccgaat tgaagcaaaa aaggatgtcg aacagctgcg 420  
 42 cgaggaatcc ggaatgtcg ttttggaatg gctgaaaacgg ttttattaaag aaaaaaaattg 480  
 44 aaaaagcggg aggagaaaaca ttg gaa aca ctt ttg gaa tta aaa aat gta tca 533  
 45 Leu Glu Thr Leu Glu Leu Lys Asn Val Ser  
 46 1 5 10  
 48 aaa acg atc agg ggg aaa aag atc atc gag ggc ttg agt ttt gac gtg 581  
 49 Lys Thr Ile Arg Gly Lys Lys Ile Ile Glu Gly Leu Ser Phe Asp Val  
 50 15 20 25  
 52 cgg gca ggc gag ata ttc ggc ttc ctg ggg ccg aac ggc gcc gga aaa 629  
 53 Arg Ala Gly Glu Ile Phe Gly Phe Leu Gly Pro Asn Gly Ala Gly Lys  
 54 30 35 40  
 56 acg acg acg atc cgg atg att gtc gga cat atg agc att acc gcc ggt 677  
 57 Thr Thr Thr Ile Arg Met Ile Val Gly His Met Ser Ile Thr Ala Gly  
 58 45 50 55  
 60 gaa atc gcc gtg tgc ggc gtc agc gta aaa gaa aac ttt gaa aag gct 725  
 61 Glu Ile Ala Val Cys Gly Val Ser Val Lys Glu Asn Phe Glu Lys Ala  
 62 60 65 70 75  
 64 gca cgg cat atc gga gcg atc gtt gaa aac acc ccg gag ctt tat aaa ttt 773

**RAW SEQUENCE LISTING**  
**PATENT APPLICATION: US/10/510,941**

**DATE: 10/22/2004**  
**TIME: 09:18:29**

**Input Set : A:\pto.kd.txt**  
**Output Set: N:\CRF4\10222004\J510941.raw**

65 Ala Arg His Ile Gly Ala Ile Val Glu Asn Pro Glu Leu Tyr Lys Phe			
66 80	85	90	
68 tta acg ggc tac caa aac ctt cag caa tac gcg cgc atg acg aaa ggc		821	
69 Leu Thr Gly Tyr Gln Asn Leu Gln Gln Tyr Ala Arg Met Thr Lys Gly			
70 95	100	105	
72 gtg acg aaa aag aaa att gac gaa atc gtc gag ctc gtc gga ttg aaa			
73 Val Thr Lys Lys Ile Asp Glu Ile Val Glu Leu Val Gly Leu Lys			
74 110	115	120	
76 aac agg atc aac gac aag gtc aaa acg tat tcg tta gga atg aga caa		917	
77 Asn Arg Ile Asn Asp Lys Val Lys Thr Tyr Ser Leu Gly Met Arg Gln			
78 125	130	135	
80 agg ctt ggc ctt gcc caa agc ctt ttg cac gat cca aag ctg ttg att		965	
81 Arg Leu Gly Leu Ala Gln Ser Leu Leu His Asp Pro Lys Leu Leu Ile			
82 140	145	150	155
84 ctc gat gag ccg acg aac ggg ctt gat ccg gca ggt atc ccg gaa atc		1013	
85 Leu Asp Glu Pro Thr Asn Gly Leu Asp Pro Ala Gly Ile Arg Glu Ile			
86 160	165	170	
88 cgt gac tat tta aga aag ctg acg aga gaa aag gga atg gcg gtc atc		1061	
89 Arg Asp Tyr Leu Arg Lys Leu Thr Arg Glu Lys Gly Met Ala Val Ile			
90 175	180	185	
92 gtt tca agc cac ctg ctt tca gaa atg gag ctg atg tgc gac agg atc		1109	
93 Val Ser Ser His Leu Leu Ser Glu Met Glu Leu Met Cys Asp Arg Ile			
94 190	195	200	
96 gcc att att caa aac gga aag ctc ccg gat att cag cat gtg cac gga		1157	
97 Ala Ile Ile Gln Asn Gly Lys Leu Arg Asp Ile Gln His Val His Gly			
98 205	210	215	
100 ccc gct ccg gat gag aag aag cgc tat tat att caa gcg gac gac acc		1205	
101 Pro Ala Arg Asp Glu Lys Lys Arg Tyr Tyr Ile Gln Ala Asp Asp Thr			
102 220	225	230	235
104 cag gct ctc aca ccg gaa gcg gct gct ttc aga aag gtg aag gtt gac		1253	
105 Gln Ala Leu Thr Arg Glu Ala Ala Ala Phe Arg Lys Val Lys Val Asp			
106 240	245	250	
108 gaa gcg gaa ggc ggg ata gag ctc agc att caa aag gat gaa gtg cct		1301	
109 Glu Ala Glu Gly Ile Glu Leu Ser Ile Gln Lys Asp Glu Val Pro			
110 255	260	265	
112 gat ttg att aaa cac ttg aca gac agc ggt gtt cgc tta tat gaa gtg		1349	
113 Asp Leu Ile Lys His Leu Thr Asp Ser Gly Val Arg Leu Tyr Glu Val			
114 270	275	280	
116 aag gct gtg aac aaa tcg ctg gaa gac cga ttc ctg gaa atc acc gca		1397	
117 Lys Ala Val Asn Lys Ser Leu Glu Asp Arg Phe Leu Glu Ile Thr Ala			
118 285	290	295	
120 gat aag gag gaa gct cag cat gtt taatctcatc gtaaaatgaat ggatcaaaat		1451	
121 Asp Lys Glu Ala Gln His Val			
122 300	305		
124 tttaatcga aaaggcacat acgtcatgat cgaaatttg ctgttagctg tcatcggt		1511	
126 gggcggtctc acaaagacga tcggagagac agacccaaac acggactgga aaaaggaatt		1571	
128 ggcgcaggaa ataaggacaa gggggcttag t		1602	
131 <210> SEQ ID NO: 2			
132 <211> LENGTH: 307			

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/10/510,941

DATE: 10/22/2004  
TIME: 09:18:29

Input Set : A:\pto.kd.txt  
Output Set: N:\CRF4\10222004\J510941.raw

```

133 <212> TYPE: PRT
134 <213> ORGANISM: Bacillus licheniformis
136 <400> SEQUENCE: 2
138 Leu Glu Thr Leu Leu Glu Leu Lys Asn Val Ser Lys Thr Ile Arg Gly
139      5          10          15
142 Lys Lys Ile Ile Glu Gly Leu Ser Phe Asp Val Arg Ala Gly Glu Ile
143      20         25          30
146 Phe Gly Phe Leu Gly Pro Asn Gly Ala Gly Lys Thr Thr Thr Ile Arg
147      35         40          45
150 Met Ile Val Gly His Met Ser Ile Thr Ala Gly Glu Ile Ala Val Cys
151      50         55          60
154 Gly Val Ser Val Lys Glu Asn Phe Glu Lys Ala Ala Arg His Ile Gly
155 65          70          75          80
158 Ala Ile Val Glu Asn Pro Glu Leu Tyr Lys Phe Leu Thr Gly Tyr Gln
159          85          90          95
162 Asn Leu Gln Gln Tyr Ala Arg Met Thr Lys Gly Val Thr Lys Lys Lys
163          100         105         110
166 Ile Asp Glu Ile Val Glu Leu Val Gly Leu Lys Asn Arg Ile Asn Asp
167          115         120         125
170 Lys Val Lys Thr Tyr Ser Leu Gly Met Arg Gln Arg Leu Gly Leu Ala
171          130         135         140
174 Gln Ser Leu Leu His Asp Pro Lys Leu Leu Ile Leu Asp Glu Pro Thr
175 145          150          155          160
178 Asn Gly Leu Asp Pro Ala Gly Ile Arg Glu Ile Arg Asp Tyr Leu Arg
179          165          170          175
182 Lys Leu Thr Arg Glu Lys Gly Met Ala Val Ile Val Ser Ser His Leu
183          180          185          190
186 Leu Ser Glu Met Glu Leu Met Cys Asp Arg Ile Ala Ile Ile Gln Asn
187          195          200          205
190 Gly Lys Leu Arg Asp Ile Gln His Val His Gly Pro Ala Arg Asp Glu
191          210          215          220
194 Lys Lys Arg Tyr Tyr Ile Gln Ala Asp Asp Thr Gln Ala Leu Thr Arg
195 225          230          235          240
198 Glu Ala Ala Ala Phe Arg Lys Val Lys Val Asp Glu Ala Glu Gly Gly
199          245          250          255
202 Ile Glu Leu Ser Ile Gln Lys Asp Glu Val Pro Asp Leu Ile Lys His
203          260          265          270
206 Leu Thr Asp Ser Gly Val Arg Leu Tyr Glu Val Lys Ala Val Asn Lys
207          275          280          285
210 Ser Leu Glu Asp Arg Phe Leu Glu Ile Thr Ala Asp Lys Glu Glu Ala
211          290          295          300
214 Gln His Val
215 305
218 <210> SEQ ID NO: 3
219 <211> LENGTH: 1938
220 <212> TYPE: DNA
221 <213> ORGANISM: Bacillus licheniformis
224 <220> FEATURE:
225 <221> NAME/KEY: CDS

```

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/10/510,941

DATE: 10/22/2004  
TIME: 09:18:29

Input Set : A:\pto.kd.txt  
Output Set: N:\CRF4\10222004\J510941.raw

226 <222> LOCATION: (501)..(1457)  
 228 <400> SEQUENCE: 3  
 229 cgcacatggac tttgaacgt cgaaaccgaa gataaaaaag gaaatgagct tgccaagcag 60  
 231 acctttcaaa aattgaagga aaccgattt aatttcatcg gcaatgtgga agcccgcat 120  
 233 atgctggacg gagtcgctga tgtcatcgac acagacggct ttaccggtaa cgttgcctt 180  
 235 aaaacggtcg agggcgccgc gctgtccatt tttaaatgc tgagaacgac gctgacttcg 240  
 237 agcttcacgg cgaagctcgc cgcttctgca ctgaagccga agctgaaaaga aatgaaaacg 300  
 239 aaaatggatt actctgaata cggcgagcc ggattgttcg gcttaaaggc gcccgcatc 360  
 241 aaagcgcacg gatcatctga cggacgcgcc gtttatcacg cgatccgcca ggccagagag 420  
 243 atggtcagcc aaaatgtcgc ggcatttatac gaagaaaaaa ttcaacaaaaa agcagatgaa 480  
 245 tagtctggag gtttaaacac atg ggc aag att gct ttt cta ttc ccg ggc caa 533  
 246 Met Gly Lys Ile Ala Phe Leu Phe Pro Gly Gln  
 247 1 5 10  
 249 ggt tcg cag cat atc ggc atg gga cac gaa ttg tat gaa aaa gaa ccg 581  
 250 Gly Ser Gln His Ile Gly Met Gly His Glu Leu Tyr Glu Lys Glu Pro  
 251 15 20 25  
 253 aat gcg aag aag att ttt gaa gaa gcg gat caa acg ctt gaa aca aaa 629  
 254 Asn Ala Lys Lys Ile Phe Glu Glu Ala Asp Gln Thr Leu Glu Thr Lys  
 255 30 35 40  
 257 ctg agc acc ctc atg ttt gaa ggg gat gca aag gaa ctg acg ctt aca 677  
 258 Leu Ser Thr Leu Met Phe Glu Gly Asp Ala Lys Glu Leu Thr Leu Thr  
 259 45 50 55  
 261 tac aac gcg cag cca agc ctt tta acg gcg agc atc gca gcg ctt gaa 725  
 262 Tyr Asn Ala Gln Pro Ser Leu Leu Thr Ala Ser Ile Ala Ala Leu Glu  
 263 60 65 70 75  
 265 aaa ctg aag gaa tac ggc att aaa gcc gac tat gcg gca ggt cac agc 773  
 266 Lys Leu Lys Glu Tyr Gly Ile Lys Ala Asp Tyr Ala Ala Gly His Ser  
 267 80 85 90  
 269 ctc ggc gaa tac agc gca ttg gtc gct gcc ggc gcc ttg tcg ttt aaa 821  
 270 Leu Gly Glu Tyr Ser Ala Leu Val Ala Ala Gly Ala Leu Ser Phe Lys  
 271 95 100 105  
 273 gat gcg gtt tat gcc gtc aga aag cgc ggc gaa ttc atg aat gaa gcc 869  
 274 Asp Ala Val Tyr Ala Val Arg Lys Arg Gly Glu Phe Met Asn Glu Ala  
 275 110 115 120  
 277 gtg ccg gcg gga gaa ggc gcg atg gcg gcc att ctc ggc atg gac agc 917  
 278 Val Pro Ala Gly Glu Gly Ala Met Ala Ala Ile Leu Gly Met Asp Ser  
 279 125 130 135  
 281 cag gcg ctg aaa gaa gtg acg gac aaa att tcc gaa gaa gga aac ctt 965  
 282 Gln Ala Leu Lys Glu Val Thr Asp Lys Ile Ser Glu Glu Gly Asn Leu  
 283 140 145 150 155  
 285 gtt cag ctc gcc aat ttg aac tgc cct ggg caa atc gtc atc tcg gga 1013  
 286 Val Gln Leu Ala Asn Leu Asn Cys Pro Gly Gln Ile Val Ile Ser Gly  
 287 160 165 170  
 289 aca gct aaa ggc gtg gag ctc gct tca gag ctt gcg aaa gaa aag ggc 1061  
 290 Thr Ala Lys Gly Val Glu Leu Ala Ser Glu Leu Ala Lys Glu Lys Gly  
 291 175 180 185  
 293 gca aaa cgc gcg att cct ctc gaa gtc agc ggg ccg ttc cat tct gag 1109  
 294 Ala Lys Arg Ala Ile Pro Leu Glu Val Ser Gly Pro Phe His Ser Glu  
 295 190 195 200

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/10/510,941

DATE: 10/22/2004  
TIME: 09:18:29

Input Set : A:\pto.kd.txt  
Output Set: N:\CRF4\10222004\J510941.raw

297	ctg	atg	aag	ccg	gca	gct	gat	aag	ctt	cgt	gaa	gtt	ctt	gat	gcg	tgc	1157
298	Leu	Met	Lys	Pro	Ala	Ala	Asp	Lys	Leu	Arg	Glu	Val	Leu	Asp	Ala	Cys	
299	205				210					215							
301	acg	atc	aac	gac	gca	gcc	att	ccg	gtc	gtc	tcc	aac	gta	acg	gcc	gac	1205
302	Thr	Ile	Asn	Asp	Ala	Ala	Ile	Pro	Val	Val	Ser	Asn	Val	Thr	Ala	Asp	
303	220				225					230				235			
305	ttt	gta	acg	gat	aaa	gac	gac	att	aag	aat	aaa	ctg	att	gaa	cag	ctg	1253
306	Phe	Val	Thr	Asp	Lys	Asp	Asp	Ile	Lys	Asn	Lys	Leu	Ile	Glu	Gln	Leu	
307					240				245			250					
309	tat	tcc	cct	gta	cgc	ttt	gaa	gaa	aca	atc	acg	cgc	ctg	att	gac	gaa	1301
310	Tyr	Ser	Pro	Val	Arg	Phe	Glu	Glu	Thr	Ile	Ser	Arg	Leu	Ile	Asp	Glu	
311					255				260			265					
313	ggc	gtc	acg	acc	ttc	att	gaa	atc	ggt	ccc	gga	aag	gtt	ttg	tca	ggg	1349
314	Gly	Val	Thr	Thr	Phe	Ile	Glu	Ile	Gly	Pro	Gly	Lys	Val	Leu	Ser	Gly	
315					270				275			280					
317	ctt	gtg	aag	aaa	gtg	aac	cgc	aga	gtc	aaa	acg	att	gct	gta	tca	gac	1397
318	Leu	Val	Lys	Lys	Val	Asn	Arg	Arg	Val	Lys	Thr	Ile	Ala	Val	Ser	Asp	
319					285				290			295					
321	ccg	aac	aca	att	gaa	ctt	gcc	gtt	caa	acg	ttg	aag	gag	gaa	aac	gaa	1445
322	Pro	Asn	Thr	Ile	Glu	Leu	Ala	Val	Gln	Thr	Leu	Lys	Glu	Glu	Asn	Glu	
323	300				305					310			315				
325	aat	gct	gga	aaa	taaaacagcc	gttgtgacag	gaggcctcaag	aggaatcgcc									1497
326	Asn	Ala	Gly	Lys													
329	cgcgcgatcg	ccctggacct	ggcgaaaaac	ggagcaaatg	tcgtcgtaa	ctacgcggga											1557
331	aatgaagcga	aagcgaacga	agtcttagac	gaaatcaaag	cgctcgcccg	cgatgcgttt											1617
333	gcttttaaag	cggaacgttcc	caatgcggat	gagggtcagg	cgatgatgaa	ggaagcggtc											1677
335	ggacgcgttcg	gcacgcgttga	catccttgta	aacaatgcgg	gcattactaa	agacaatctg											1737
337	ttcatgagaa	tgaaaagaaga	tgaatgggac	gacgtcatta	acataaactt	aaaagggtgt											1797
339	ttcaattgtt	caaaaagctgt	gacaagacag	atgatgaaac	aaagaagcgg	ccggatcatc											1857
341	aatatcacct	cggtttagg	cgtcgctgg	aacgcggg	aggccaacta	tgtcgccgct											1917
343	aaatcaggcg	tattccagta	g														1938
346	<210>	SEQ ID NO:	4														
347	<211>	LENGTH:	319														
348	<212>	TYPE:	PRT														
349	<213>	ORGANISM:	Bacillus licheniformis														
351	<400>	SEQUENCE:	4														
353	Met	Gly	Lys	Ile	Ala	Phe	Leu	Phe	Pro	Gly	Gln	Gly	Ser	Gln	His	Ile	
354	1				5				10				15				
357	Gly	Met	Gly	His	Glu	Leu	Tyr	Glu	Lys	Glu	Pro	Asn	Ala	Lys	Lys	Ile	
358					20				25			30					
361	Phe	Glu	Glu	Ala	Asp	Gln	Thr	Leu	Glu	Thr	Lys	Leu	Ser	Thr	Leu	Met	
362					35				40			45					
365	Phe	Glu	Gly	Asp	Ala	Lys	Glu	Leu	Thr	Leu	Thr	Tyr	Asn	Ala	Gln	Pro	
366					50				55			60					
369	Ser	Leu	Leu	Thr	Ala	Ser	Ile	Ala	Ala	Leu	Glu	Lys	Leu	Lys	Glu	Tyr	
370	65				70					75			80				
373	Gly	Ile	Lys	Ala	Asp	Tyr	Ala	Ala	Gly	His	Ser	Leu	Gly	Glu	Tyr	Ser	
374					85				90			95					
377	Ala	Leu	Val	Ala	Ala	Gly	Ala	Leu	Ser	Phe	Lys	Asp	Ala	Val	Tyr	Ala	

**VERIFICATION SUMMARY**

PATENT APPLICATION: US/10/510,941

DATE: 10/22/2004

TIME: 09:18:30

Input Set : A:\pto.kd.txt

Output Set: N:\CRF4\10222004\J510941.raw

L:13 M:270 C: Current Application Number differs, Replaced Current Application No

L:13 M:271 C: Current Filing Date differs, Replaced Current Filing Date